PERGAMON

Journal of Structural Geology 26 (2004) 143–153

JOURNAL OF STRUCTURAL GEOLOGY

www.elsevier.com/locate/jsg

Fitting an ellipse to an arbitrary shape: implications for strain analysis

Kieran F. Mulchrone[a,*], Kingshuk Roy Choudhury[b]

[a]*Department of Applied Mathematics, National University of Ireland, Cork, Republic of Ireland*
[b]*Department of Statistics, National University of Ireland, Cork, Republic of Ireland*

Received 29 May 2002; received in revised form 12 May 2003; accepted 15 May 2003

Abstract

An ellipse can be fit to an arbitrary shape using a linear least squares approach applied to boundary data. Alternatively, this problem can also be solved by calculating the second moments of the entire region, a technique popular in image analysis applications. If the irregular shape can be approximated by a polygon then Greens theorem allows efficient calculation of the second moments. If the shape is pixelated then the second moments can be calculated by a simple summation process. By considering the behaviour of these fitting methods with increasing deformation it is shown that as an arbitrary shape passively deforms, the best-fit ellipse also behaves as if it were deforming passively. This implies that all techniques of strain analysis that were previously restricted to populations of elliptical objects may now be applied to populations of arbitrary shapes, provided the best-fit ellipse is calculated by one of the methods described here. Furthermore it implies that selective sampling based on shape or methods of weighting based upon shape are invalid and tend to bias the raw data.

Keywords: Ellipse; Fitting; Strain analysis; Moments; Weighting

1. Introduction

A large proportion of methods for strain analysis require that the strain markers used are of approximately elliptical shape, deform passively and that marker long axes were initially uniformly distributed (Ramsay, 1967, pp. 202–209; Dunnet, 1969; Elliott, 1970; Matthews et al., 1974; Shimamoto and Ikeda, 1976; Lisle, 1977a, 1994; Mulchrone and Meere, 2001; Mulchrone et al., 2003). However, in reality, most natural strain markers are only poorly approximated by an ellipse and therefore the accuracy of strain estimations using ellipse-based methods may be called into question. The primary objective of this paper is to investigate the effect of ellipse fitting on these ellipse-based methods. The results of these considerations are of fundamental importance to strain analysis and its applicability.

Methods of strain analysis based on the assumption of passive behaviour of populations of strain markers remain popular and have been applied in a broad range of geological contexts over the years. For example, they have been applied to oolitic limestones (Dunnet, 1969; Elliott, 1970), pisolitic tuffs (Dunnet and Siddans, 1971; Matthews et al., 1974), lapilli tuffs (Borradaile, 1987), polymictic grits (Dunnet, 1969), modern and ancient pebble conglomerates (Lisle, 1977b; Borradaile, 1987) and grains in modern and ancient clastic sandstones (Paterson and Yu, 1994; Meere, 1995; Bresser and Walter, 1999; Mulchrone and Meere, 2001; Mulchrone et al., 2003). More recently Zulauf et al. (2002) applied the R_f/ϕ technique to determine finite strain using pebbles in metaconglomerates and ooids in slates. Bhattacharyya and Hudleston (2001) used the shapes of irregularly shaped feldspar aggregates to quantify strain in shear zones, whereas Mazzoli et al. (2001) applied the R_f/ϕ technique to deformed sub-elliptical radiolarian microfossils in limestone. Simancas et al. (2000) studied the deformation of granites using sometimes irregularly shaped quartz micrograin aggregates. It is important to know whether the inherent assumption of strain marker ellipticity has any bearing on the accuracy and applicability of these commonly used methods.

There are also a smaller number of methods that do not make any assumptions about shape. For example the method of Robin (1977) assumes no preferred orientation of the initial strain marker distribution, but requires an independent estimate of the finite strain ellipse orientation.

* Corresponding author. Tel.: +353-21-4903411; fax: +353-21-4270813.
E-mail address: k.mulchrone@ucc.ie (K.F. Mulchrone).

0191-8141/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0191-8141(03)00093-2

Mulchrone et al. (2003) have recently demonstrated that for simulated data, the mean radial length of strain markers accurately estimates the finite strain ellipse principal extension axis. Therefore the requirement of an independent estimate of the strain ellipse orientation is not very prohibitive if the strain marker long axes are easily identified. Fry-like strain analysis methods (Fry, 1979; Erslev, 1988; Erslev and Ge, 1990; McNaught, 1994) and the Delaunay triangulation nearest neighbour method (DTNNM) of Mulchrone (2003) do not require any particular shape for strain markers but do assume that strain marker centroids are anti-clustered before deformation and move relative to each other in a passive manner during deformation.

2. Fitting an ellipse to an arbitrarily shaped region

2.1. Introduction

Fitting an ellipse to an arbitrarily shaped region has received minor attention in the geological literature but is a fundamental problem for automated image analysis and computer vision techniques and has been studied in considerable detail. There are two approaches to ellipse fitting: methods based on (1) the boundary and (2) the region. For regions that are fairly regular (i.e. almost elliptical in shape) the two approaches are indistinguishable; however, for irregular regions the results are different (see Fig. 1). Region-based methods are affected less by boundary irregularities (i.e. are more robust), whereas boundary-based methods are strongly affected by irregularities. Furthermore, complex regions, such as those with one or more holes, are handled without difficulty by region-based methods, whereas most boundary-based methods inherently assume a single, outer boundary and do not consider the possibility of internal boundaries. Because strain analysis methods are generally concerned with shapes and how shapes change during deformation, the more robust region-based method is preferred here. For a review of the ellipse equation and parameters see the appendices.

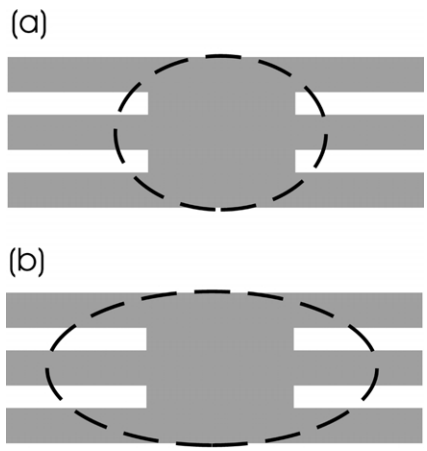

Fig. 1. Schematic illustration of the difference between region and boundary-based ellipse fitting methods after Russ (1999, p. 553). (a) Methods based on the moments of the entire region tend to be less affected by irregular boundaries. (b) Methods based on the boundary shape are disproportionally influenced by boundary irregularities.

2.2. Boundary-based methods

Boundary methods consider that the arbitrary region consists of a set of points sampled from the region boundary. For example Hart and Rudman (1997) developed a linear least squares fitting for the Cartesian equation of an ellipse centred on the origin and a similar technique was used by Erslev and Ge (1990). Workers in image analysis and computer vision have employed a variety of techniques including linear least squares, weighted least squares, Kalman filtering and robust estimation methods (Zhang, 1997). Fitzgibbon et al. (1999) presented a direct ellipse-specific fitting technique, which minimised the algebraic distance subject to an ellipse-specific constraint and has been greatly improved by Halir and Flusser (1998). All of the mentioned boundary-based methods have been successfully applied in practice, although the method of Fitzgibbon et al. (1999) has distinct advantages in that it always produces an ellipse no matter how bad the input data.

Hart and Rudman (1997) used linear least squares to fit the Cartesian equation of an ellipse centred on the origin (data not naturally centred on the origin needs to be transformed such that the centroid of the shape is located at the origin where the centroid is calculated from the moments as shown in the next section and in the appendix):

$$ax^2 + bxy + cy^2 = \rho^2 \tag{1}$$

where a, b, c and ρ are unknown parameters, so that for a set of points (x_i,y_i) where i goes from one to n, the total number of points, we wish to minimise:

$$\sum_{i=1}^{n}\left(ax_i^2 + bx_iy_i + cy_i^2 - \rho^2\right)^2 \tag{2}$$

Alternatively the method of Fitzgibbon et al. (1999) or Halir and Flusser (1998) can be employed for the general equation of a conic (not necessarily centred on the origin):

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{3}$$

where a, b, c, d, e and f are unknown parameters, subject to the ellipse-specific constraint:

$$b^2 < -4ac \tag{4}$$

which ensures that the curve is an ellipse and not a hyperbola or parabola. Introducing the notation:

$$S_{x^\alpha y^\beta} = \sum_{i=1}^{n} x_i^\alpha y_i^\beta \tag{5}$$

where α and $\beta$are integers, then for both methods the least

squares solution depends only on the following:

$$S_1,\ S_x,\ S_y,\ S_{x^2},\ S_{y^2},\ S_{xy},\ S_{x^3},\ S_{y^3},\ S_{x^2y},\ S_{xy^2},\ S_{x^4},\ S_{y^4},\ \times S_{x^2y^2},\ S_{xy^3},\ S_{x^3y} \tag{6}$$

where $S_1 = S_{x^0y^0} = n$. For exact details of the solutions, the reader should refer to the original publications.

2.3. Region-based methods

Region-based methods use the moments of a shape in estimating the best-fit ellipse and are extremely popular in image analysis applications (Jain, 1989, pp. 392–394; Jähne, 1997, pp. 509–512; Russ, 1999, p. 552). The best-fit ellipse is used as a descriptive measure of shape even though usually the shape is far from elliptical. As this method is favoured here, it is described in detail (see Gonzalez and Wintz, 1987, pp. 419–423). The moment of order $p + q$ (where p and q are integers) of a 2D region (G) is calculated by the following integral evaluated over the area of G:

$$m_{pq} = \int\int_G f(x,y)x^p y^q \mathrm{d}x\mathrm{d}y \tag{7}$$

The function f is used to describe situations where some property varies around the area (such as physical density) and in the present case no such variation is present because all parts of the region are equally important, so that $f(x,y) = 1$. Physically, the zeroth moment is equal to the area of G while $\left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right)$ are the centroids of G. The central moments are given by:

$$u_{pq} = \int\int_G f(x-\bar{x})^p (y-\bar{y})^q \mathrm{d}x\mathrm{d}y \tag{8}$$

where:

$$\bar{x} = \frac{m_{10}}{m_{00}} \tag{9}$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \tag{10}$$

and can be written in terms of the moments, for example:

$$u_{00} = m_{00} \tag{11}$$

$$u_{10} = u_{01} = 0 \tag{12}$$

$$u_{20} = m_{20} - \frac{m_{10}^2}{m_{00}} \tag{13}$$

$$u_{02} = m_{02} - \frac{m_{01}^2}{m_{00}} \tag{14}$$

$$u_{11} = m_{11} - \frac{m_{10}m_{01}}{m_{00}} \tag{15}$$

and the normalised central moments are defined as:

$$n_{pq} = \frac{u_{pq}}{u_{00}^{\gamma}} \tag{16}$$

where

$$\gamma = \frac{p+q}{2} + 1 \tag{17}$$

and utilised in some applications.

The central moments are then normally used to define the best-fit ellipse (Hu, 1962; Teague, 1980) for the arbitrary shape with semi-minor and major axes a and b, respectively, and long axes oriented at ϕ:

$$\Delta = \sqrt{4u_{11}^2 + (u_{20} - u_{02})^2} \tag{18}$$

$$\phi = \frac{1}{2}\tan^{-1}\left(\frac{2u_{11}}{u_{20} - u_{02}}\right) \tag{19}$$

$$a = \sqrt{\frac{2(u_{20} + u_{02} + \Delta)}{u_{11}}} \tag{20}$$

$$b = \sqrt{\frac{2(u_{20} + u_{02} - \Delta)}{u_{11}}} \tag{21}$$

A derivation of these equations is included in the appendices. Most methods of strain analysis involving elliptical strain markers are not concerned with the absolute values of a and b but with the axial ratio R:

$$R = \sqrt{\frac{u_{20} + u_{02} + \Delta}{u_{20} + u_{02} - \Delta}} \tag{22}$$

This method of ellipse fitting essentially equalises the second moments of the shape and the ellipse (Jain, 1989, p. 394).

If the normalised central moments of an arbitrary shape can be calculated then the best fitting ellipse may be estimated using Eqs. (18)–(22). In reality, evaluating integrals such as Eqs. (7) and (8) is not practicable, so that typically arbitrary shapes are approximated by either a polygonal representation of the boundary or as a pixelated digital image. The moments of both representations are easily calculated and details are presented in the appendices.

3. Relationship between an arbitrary shape and its fitted ellipse during deformation

Methods of strain analysis based on elliptical strain markers make several fundamental assumptions:

1. Strain markers are approximately elliptically shaped.
2. Deformation is homogeneous over the region of the strain markers.

3. The viscosity contrast between markers and matrix is negligible leading to passive behaviour.

Hence in what follows it is assumed that points and curves behave passively during deformation. A pure shear deformation is considered whereby the long axis of the finite strain ellipse is parallel to the Cartesian x-axis. This does not affect the generality of the result due to the polar decomposition theorem, which proves that any general deformation is composed of a pure shear deformation and a rotation (Mulchrone, 2002) and because all points and curves behave passively the component of rotation does not affect their relative distribution.

During this pure shear deformation points are modified as follows:

$$x' = \sqrt{R_s}x \tag{23}$$

$$y' = \frac{y}{\sqrt{R_s}} \tag{24}$$

where (x,y) is the point in the undeformed state and (x',y') is the point in the deformed state and R_s is the axial ratio of the strain ellipse. A region G in the undeformed state is transformed into the region G' in the deformed state. Hence the moments of the shape in the deformed state are:

$$m'_{pq} = \int\int_{G'} x'^p y'^q \mathrm{d}x'\mathrm{d}y' \tag{25}$$

$$= \frac{\sqrt{R_s^p}}{\sqrt{R_s^q}} \int\int_{G'} x^p y^q \mathrm{d}x\mathrm{d}y \tag{26}$$

$$= R_s^{\frac{p-q}{2}} m_{pq} \tag{27}$$

Clearly the same transformation rule applies to both the central and normalised central moments, i.e.:

$$u'_{pq} = R_s^{\frac{p-q}{2}} u_{pq} \tag{28}$$

$$n'_{pq} = R_s^{\frac{p-q}{2}} n_{pq} \tag{29}$$

Eqs. (18)–(22) describe the best-fit ellipse to an arbitrary shape but can also be used to describe an ellipse in terms of moments. Clearly then application of the transformation in Eq. (19) to Eqs. (18)–(22) describes how a, b, ϕ and R change during deformation, i.e. these are equations for the passive behaviour of an ellipse during deformation. Therefore as an arbitrary shape passively deforms, the best-fit ellipse also behaves as if it were passively deforming (see Fig. 2). This is an important result with implications for strain analysis.

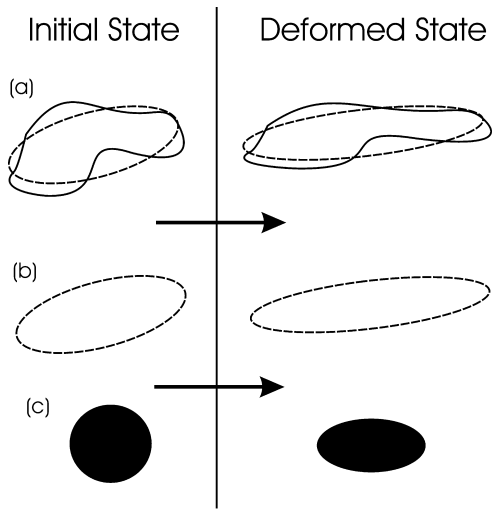

Fig. 2. (a) An arbitrary shape with an ellipse fitted in the undeformed state. After deformation the shape has changed and an ellipse is fitted to the deformed shape. (b) If the ellipse, which was fit in the undeformed state, was passively deformed it would coincide with the ellipse fit in the deformed state, due to the way moments behave during passive deformation. (c) The finite strain ellipse for the deformation.

Furthermore, a similar argument applies to best-fit ellipses calculated using boundary-based methods as the best-fit ellipses depend on the values given in Eq. (6), which also transform during deformation as follows:

$$S'_{x^\alpha y^\beta} = R_s^{\frac{a-b}{2}} S_{x^\alpha y^\beta} \tag{30}$$

4. Implications for strain analysis and a simulation study

The first implication is that existing methods for strain analysis using elliptical strain markers can be applied to materials containing assemblages of arbitrary shapes provided that the ellipse parameters are measured using either the moment-based best-fit ellipse method or the linear least-squares boundary methods. Although this conclusion may also be applied to other ellipse-fitting techniques, these have not been checked. The basic assumptions of these strain analysis techniques must still apply, so that in the undeformed state the best-fit ellipse long-axis orientations should be uniformly distributed and the axial ratio distributions should be independent of orientation (Mulchrone et al., 2003). Previously, only the method of Robin (1977) could be applied to arbitrary shapes.

Secondly, this result implies that data acquisition strategies that tend to ignore non-elliptical objects (i.e. in an attempt to avoid restrictions inherent in the methodologies themselves) actually tends to distort the results of the analysis and produces a less accurate strain estimate by artificially concentrating on a subset of the object population. In other words this sampling approach introduces a bias into the dataset for no good reason. This

implication has been empirically verified by a simulation study presented in brief in the next section, although the effect of the bias is small.

The third implication is related to the previous one and is that the accuracy of a strain estimate cannot be improved by weighting the data according to how close an object is to an ellipse. In fact by making such a hypothesis the author discovered the result presented in this paper. The idea was that by weighting the input data according to how good it fitted to an ellipse, the strain estimate could be improved upon. In actuality the converse is true as demonstrated by the simulation study.

There are also practical implications to be considered. As far as the author is aware, there is no published consensus on how best to estimate ellipse parameters for strain markers. In the author's experience some people concentrate on visually finding the longest line through a shape and then take the longest length perpendicular to this direction (often referred to as a callipers approach), whereas others may try to imagine the best-fit ellipse and calculate the long and short-axis of this mental ellipse. The result is that no two geologists would ever produce the same set of data from a particular sample. Using one of the ellipse-fitting methods described above eliminates this type of subjectivity and would allow for valid comparisons of results between workers. However, the human fitting methods have the advantage of being applicable in almost any situation, either in the field or under the microscope. It is not feasible to calculate the best-fit ellipse using the quantitative methods described above in the field or under the microscope. The ultimate resolution of this problem will come from digital image analysis techniques, which seek to automatically identify strain markers from field photographs or micro-photographs (Ailleres and Champenois, 1994; Ailleres et al., 1995; Heilbronner, 2000) combined with the quantitative fitting techniques described here.

In order to empirically demonstrate the theoretical considerations presented above a small simulation study was carried out. Software was written in Microsoft Visual C++ version 6.0, which does the following:

1. Generates a population of non-intersecting polygons whose long axis orientations are uniformly distributed (see Fig. 3a for an example).
2. Progressively strain the population (see Fig. 3b for an example of a strained population).
3. For each imposed strain value (R_{sact}) the best-fit ellipse is calculated for each polygonal object.
4. An estimate of the strain ellipse (R_{scalc}) along with associated confidence intervals is calculated using the bootstrap (Mulchrone et al., 2003) firstly by application of a weighting according to goodness of fit (the better the fit the higher the weight) and secondly by application of no weighting at all.

The applied weighting should reflect how close a particular shape is to its associated best-fit ellipse. Highly irregular shapes should be given less weighting than those that closely approximate an ellipse. To comply with these requirements, the weighting is calculated as the ratio of total shape area to the area of the overlap between the shape and its best-fit ellipse. Therefore, if the overlap area is small (i.e. the shape closely approximates the best-fit ellipse) the weighting is large and vice versa. The values of the weights may be normalised to lie between zero and one, although this is unnecessary in the case of the method of Mulchrone et al. (2003) due to the particular form of the equations used.

A small modification needs to be made to the method of strain analysis devised by Mulchrone et al. (2003) in order to incorporate weighting. The modified method involves the following calculations for a population of n ellipses with axial ratios R_i, orientations ϕ_i and weights w_i, calculated as described above. For each ellipse the following values are calculated:

$$m_i = \frac{1}{2}\left(R_i - \frac{1}{R_i}\right) \tag{31}$$

$$p_i = \frac{1}{2}\left(R_i + \frac{1}{R_i}\right) \tag{32}$$

subsequently, the following averages may be determined:

$$q_{\mathrm{s}} = \frac{1}{n}\sum_{i=1}^{n} w_i p_i \tag{33}$$

$$t_{\mathrm{s}} = \frac{1}{n}\sum_{i=1}^{n} w_i m_i \cos(2\phi_i) \tag{34}$$

$$u_{\mathrm{s}} = \frac{1}{n}\sum_{i=1}^{n} w_i m_i \sin(2\phi_i) \tag{35}$$

and the strain ellipse parameters (axial ratio R_{s} and orientation ϕ_{s}) calculated from:

$$\tan(2\phi_{\mathrm{s}}) = \frac{u_{\mathrm{s}}}{t_{\mathrm{s}}} \tag{36}$$

$$R_{\mathrm{s}} = \sqrt{\frac{q_{\mathrm{s}}\cos(2\phi_{\mathrm{s}}) + t_{\mathrm{s}}}{q_{\mathrm{s}}\cos(2\phi_{\mathrm{s}}) - t_{\mathrm{s}}}} \tag{37}$$

In the original method, weighting was not considered and that was equivalent to letting each $w_i = 1$ in the modified method.

The results of the simulation study are presented in graphical form in Fig. 4 for a population of 100 polygons. Errors tend to increase with increasing applied tectonic strain (R_{sact}). The use of weighting does not decrease the size of the error and in many cases the error increases slightly by comparison with the unweighted data; however, this effect is quite small in relation to the size of the overall error. This is to be expected as the weighting is artificial and introduces an unwarranted bias to the data.

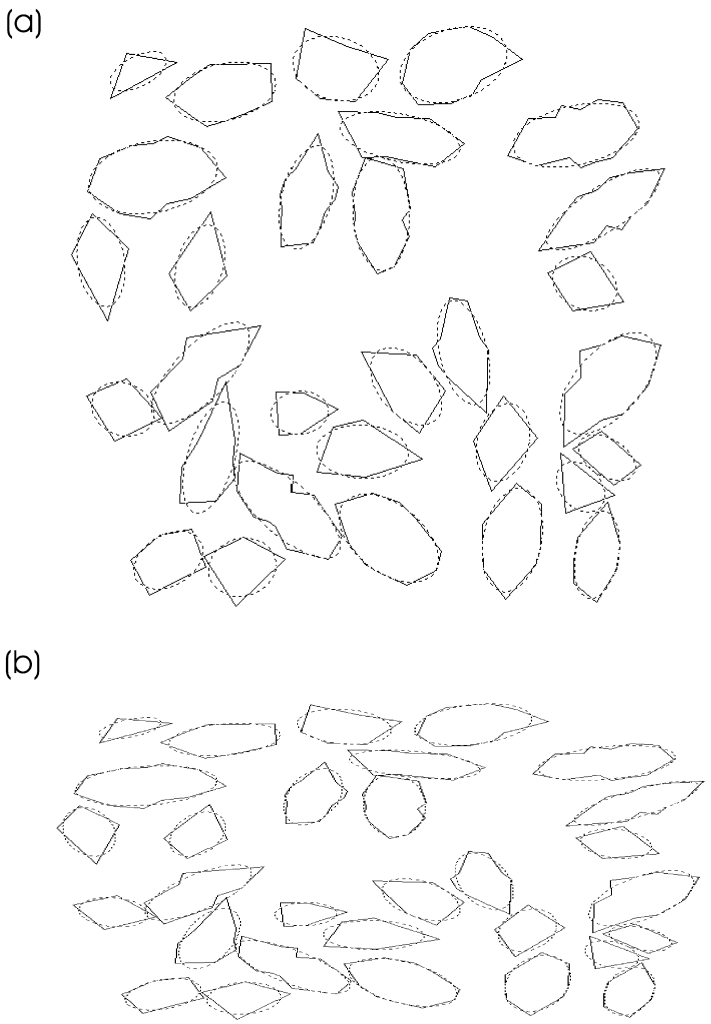

Fig. 3. (a) A set of arbitrarily shaped polygons and their best-fit ellipses in the undeformed state. (b) The same set after deformation (note that the polygons were re-scaled to allow the diagram to fit on the page).

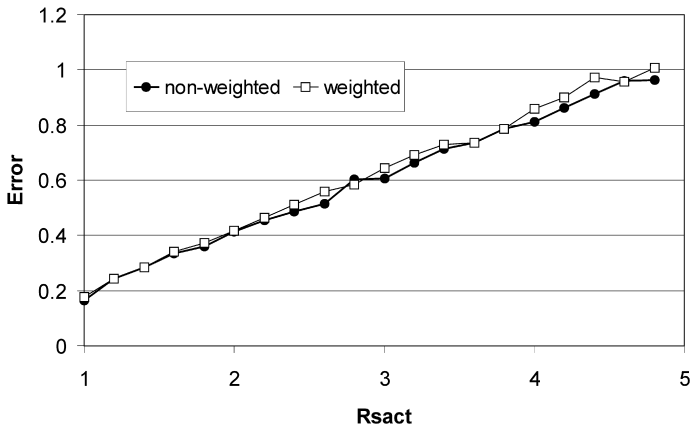

Fig. 4. A plot of error associated with calculating the axial ratio of the finite strain ellipse using the method of Mulchrone et al. (2003) against the imposed finite strain axial ratio. Results for weighted and non-weighted data are illustrated and the application of a weighting based on goodness of fit tends to slightly decrease the accuracy of the calculation.

5. Conclusions

An ellipse can be fit to an arbitrary shape using a method based on the second moments of the shape. By considering the behaviour of the moments with strain it has been shown that as an arbitrary shape passively deforms, the best-fit ellipse also behaves as if it were deforming passively. This implies that all techniques of strain analysis that were previously restricted to populations of elliptical objects may now be applied to populations of arbitrary shapes, provided the best-fit ellipse is calculated by one of the methods described above. Furthermore it implies that selective sampling or methods of weighting based upon shape are invalid and tend to bias the raw data.

Acknowledgements

The final form of the paper was improved by reviews from Dr S.K. Samanta and Dr K. Kanagawa.

Appendix A

A.1. Calculating the moments of an arbitrary polygon

This section describes how to calculate the moments of an arbitrary polygon using Green's Theorem (Yang and Albregtsen, 1996; Turkowski, 1997).

A.1.1. Parameterised representation of a line segment

It is convenient to represent the line segments that make up the polygon using a parameterised formulation as this simplifies some of the later calculations. Let our polygon be defined by a set of points $p_i = (x_i, y_i)$ where i goes from one to n, the total number of points and that the points are ordered counter-clockwise around the polygon. The polygon is composed of a set of line segments with end points (p_i,p_j), where $j = 1 + \%n$ and $\%$ is the modulus operator. The parametric equation of each line segment is:

$$x(t) = x_i + t(x_j - x_i)$$

$$y(t) = y_i + t(y_j - y_i)$$

where t is a parameter which varies from zero to one. Additionally this means that:

$$\mathrm{d}x = (x_j - x_i)\mathrm{d}t$$

$$\mathrm{d}y = (y_j - y_i)\mathrm{d}t$$

A.1.2. Green's Theorem

Green's Theorem is a well-known result from vector calculus. Let G be a bounded, closed region in the plane, whose boundary is denoted by ∂G and G can be decomposed into a finite number of bounded, closed, simple regions. Given two continuously differentiable functions $P(x,y)$ and $Q(x,y)$ on G then Green's Theorem states that:

$$\int\int_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y = \oint_{\partial G} \left[P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y\right]$$

This theorem allows us to replace an integral over the entire region G with an integral around the boundary ∂G, which is usually much simpler to calculate.

A.1.3. The moments of a polygonal region

General formulation: In general the moments of a region G are defined as:

$$m_{pq} = \int\int_G F(x,y)\mathrm{d}x\mathrm{d}y$$

where $F(x,y) = x^p y^q$ in the present case and is continuously differentiable on G. In order to apply Green's Theorem in evaluating the moment integral the following equation must hold:

$$F(x,y) = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$$

If the form of F permits it, a variety of equivalent expressions for the moment can be derived by choosing different functional values for Q and P. For example, by letting $P(x,y) = 0$ then:

$$F(x,y) = \frac{\partial Q}{\partial x}$$

and applying Green's Theorem:

$$m_{pq} = \oint_{\partial G} \left[\int F(x,y)\mathrm{d}x\right]\mathrm{d}y$$

We could also formulate this by letting $Q(x,y) = 0$ and $F(x,y) = -\frac{\partial P}{\partial x}$ or alternatively letting $F(x,y)$ be divided between $\frac{\partial Q}{\partial x}$ and $\frac{\partial P}{\partial y}$ if possible. Hence we could derive many different but equivalent formulations for each moment. The formulas presented below are the shortest version in each case.

Zeroth moments: The zeroth moment is more commonly

known as the area and is denoted by m_{00} as:

$$m_{00} = \int\int_G \mathrm{d}x\mathrm{d}y$$

where $F(x,y) = 1 = x^0y^0$ and applying Green's Theorem:

$$m_{00} = \oint_{\partial G} x\mathrm{d}y$$

For a single parameterised line segment this becomes:

$$= \int_0^1 \Big(x_i + t\big(x_j - x_i\big)\Big)\big(y_j - y_i\big)\mathrm{d}t = -\frac{1}{2}\big(x_i - x_j\big)\big(y_i - y_j\big)$$

so that summing for every line segment in the polygon, the area is:

$$m_{00} = \sum_{i=1}^{n} -\frac{1}{2}\big(x_i + x_j\big)\big(y_i - y_j\big)$$

For the case of the zeroth moment we could also split $F(x,y)$ such that $\frac{\partial Q}{\partial x} = \frac{1}{2}$ and $\frac{\partial P}{\partial y} = -\frac{1}{2}$, giving the equivalent, but more succinct:

$$m_{00} = \sum_{i=1}^{n} \frac{1}{2}\big(x_i y_j - x_j y_i\big)$$

Higher order moments: The first moments give the centroid of the polygon and are:

$$m_{10} = \int\int_G x\mathrm{d}x\mathrm{d}y$$

$$m_{01} = \int\int_G y\mathrm{d}x\mathrm{d}y$$

and applying the above analysis so that for a polygon we have:

$$m_{10} = \sum_{i=1}^{n} -\frac{1}{6}\big(x_i^2 + x_i x_j + x_j^2\big)\big(y_i - y_j\big)$$

$$m_{01} = \sum_{i=1}^{n} \frac{1}{6}\big(x_i - x_j\big)\big(y_i^2 + y_i y_j + y_j^2\big)$$

The second moments are related to inertia and are given by:

$$m_{20} = \int\int_G x^2\mathrm{d}x\mathrm{d}y$$

$$m_{02} = \int\int_G y^2\mathrm{d}x\mathrm{d}y$$

$$m_{11} = \int\int_G xy\mathrm{d}x\mathrm{d}y$$

and for a polygon may be calculated by:

$$m_{20} = \sum_{i=1}^{n} -\frac{1}{12}\big(x_i^3 + x_i^2 x_j + x_i x_j^2 + x_j^3\big)\big(y_i - y_j\big)$$

$$m_{02} = \sum_{i=1}^{n} \frac{1}{12}\big(x_i - x_j\big)\big(y_i^3 + y_i^2 y_j + y_i y_j^2 + y_j^3\big)$$

$$m_{11} = \sum_{i=1}^{n} \frac{1}{24}\big(x_i - x_j\big) \times \Big(x_i\big(3y_i^2 + 2y_i y_j + y_j^2\big) + x_j\big(y_i^2 + 2y_i y_j + 3y_j^2\big)\Big)$$

In practice we can optimise calculation of the moments by first calculating $x_i^3, x_i^2, y_i^3, y_i^2, x_j^3, x_j^2, y_j^3, y_j^2, x_i^2x_j, x_ix_j^2, y_i^2y_j$, $y_iy_j^2$, etc. so that these quantities are not evaluated more than once. Clearly the central and normalised central moments may be calculated from m_{pq}.

A.2. Calculating the moments of a pixelated region

This is a considerably simpler calculation but may be more resource intensive. Suppose that the shape or region for which we wish to calculate the moments has been pixelated in a binary digital image (i.e. if a pixel is in the shape it has value one, otherwise it has value zero). More formally suppose the region of interest is totally enclosed in a rectangular region G of size n by m pixels. Let the function $b(i,j)$, where i goes from one to n and j goes from one to m, have value one if the i,jth pixel is in the shape and zero otherwise. Then the moment calculation reduces to:

$$m_{pq} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^p y_j^q b(i,j)$$

where (x_i,y_j) is the coordinate of the i,jth pixel.

Appendix B

B.1. General equation of an ellipse in two dimensions

From basic geometry the equation of the locus of a point $\mathbf{p}' = (x',y')$ on the boundary of an ellipse, with centre $\boldsymbol{\mu} = (\mu_x', \mu_y')$, l_1 and l_2 being the length of its major and minor axes, respectively, and major axis parallel to the x'-axis, is given by:

$$\frac{\big(x' - \mu_x'\big)^2}{l_1^2} + \frac{\big(y' - \mu_y'\big)^2}{l_2^2} = 1 \tag{38}$$

In general the major axis makes at an arbitrary angle ϕ with the x'-axis and the equation of this more general ellipse can be obtained from Eq. (38) by applying a rotation of the co-ordinates of the point $\mathbf{p}' = (x',y')$ through an angle ϕ. The rotation produces new variables $\mathbf{p} = (x,y)$ as follows:

$$x = x'\cos\phi - y'\sin\phi \tag{39}$$

$$y = x'\sin\phi - y'\cos\phi \tag{40}$$

or in vector-matrix notation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{41}$$

Introducing this change of variables into Eq. (38) produces a much bigger expression, which is often represented by a general quadratic form in two variables:

$$ax^2 + by^2 + cxy + dx + ey + f = 0 \tag{42}$$

Eq. (42) is the form that is commonly used for fitting an ellipse to points using boundary based methods, as outlined in Section 2.2. However, for region based fitting, it is more convenient to represent the equation in a vector-matrix format. Notice that Eq. (38) can be written in this format as follows:

$$\begin{pmatrix} x' - \mu'_x & y' - \mu'_y \end{pmatrix} \begin{pmatrix} l_1^2 & 0 \\ 0 & l_2^2 \end{pmatrix} \begin{pmatrix} x' - \mu'_x \\ y' - \mu'_y \end{pmatrix} = 1 \tag{43}$$

The change of variables can be introduced using the matrix form of the rotation transformation in Eq. (41), yielding:

$$\begin{aligned} &\begin{pmatrix} x - \mu_x & y - \mu_y \end{pmatrix} \\ &\times \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}^{-1} \begin{pmatrix} l_1^2 & 0 \\ 0 & l_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}^{-1} \\ &\times \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} = 1 \end{aligned} \tag{44}$$

where $\boldsymbol{\mu} = (\mu_x,\mu_y)$ are the rotated coordinates of the ellipse centre. This is the matrix form of Eq. (42). If we designate the central matrix as Ξ^{-1}, i.e.

$$\Xi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} l_1^2 & 0 \\ 0 & l_2^2 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \tag{45}$$

We get the general equation of an ellipse in two dimensions (in matrix notation) as:

$$(\mathbf{p} - \boldsymbol{\mu})^{\mathrm{T}} \Xi^{-1} (\mathbf{p} - \boldsymbol{\mu}) = 1 \tag{46}$$

Here $(\mathbf{p} - \boldsymbol{\mu})^{\mathrm{T}} = (x - \mu_x \; y - \mu_y)$ and T denotes matrix transpose.

B.2. Fitting an ellipse to a region identified by points

In this section a method for fitting an ellipse E to a given a region R containing n points $\mathbf{p}_i = (x_i,y_i)$, $i = 1,2,\ldots,n$, is derived. An objective way of doing this is to define a numerical criterion which measures how well, or how closely, the ellipse E fits the region R, or equivalently, its points. The best fitting ellipse E^* is the one that minimises this criterion. The closeness of a point to an ellipse E (as represented in Eq. (46)) can be measured in terms of the squared Euclidean distance between the point and the centre of the ellipse, namely:

$$d(\mathbf{p}_i, E) = (x_i - \mu_x)^2 + \left(y_i - \mu_y\right)^2 \tag{47}$$

This distance is minimised when the centre of the ellipse E is the point itself. However, when there are many points in the region to consider, it is impossible to minimise individual distances. An aggregate distance must be minimised and an obvious choice for this is the sum of the individual distances across the different points, i.e.:

$$d_1(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E) = \sum_{i=1}^{n} \left[(x_i - \mu_x)^2 + \left(y_i - \mu_y\right)^2 \right] \tag{48}$$

Criterion $d_1(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$ can be viewed as a 'least squares' approach to fitting the ellipse. Least squares methods have been widely used to fit curves to points since the time of Gauss in the 18th century. In many cases, the solutions are simple to implement and their quality of fit is also widely understood. In the case of $d_1(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$, one can show with some simple differential calculus that the centre of the best fitting ellipse E_* will be the centroid of the points $\bar{\mathbf{p}} = (\bar{x}, \bar{y})$, where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \text{ and } \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{49}$$

are the averages across the co-ordinates of the points in the region R. However, minimising the distance does not yield any information about other parameters of the best fitting ellipse, namely l_1, l_2 and ϕ. This is because the distance does not involve these parameters in measuring closeness, only the centre. This suggests formulating a distance that involves all these parameters.

One candidate that satisfies this requirement is the following:

$$d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E) = \sum_{i=1}^{n} (\mathbf{p}_i - \boldsymbol{\mu})^{\mathrm{T}} \Xi^{-1} (\mathbf{p}_i - \boldsymbol{\mu}) \tag{50}$$

Criterion $d_1(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$ can be viewed as a variant of $d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$ with $\Xi = \mathbf{I}$, the identity matrix. Geometrically, $d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$ can be viewed as the Euclidean distance computed on variables transformed in such a way that the ellipse would appear as a circle in the transformed space. From an ellipse fitting viewpoint, Eq. (45) tells us that Ξ incorporates l_1, l_2 and ϕ. Therefore it should be

possible to recover information about these parameters (in addition to the centre) by minimising $d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$. However, minimising criterion $d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$ is not as mathematically simple as minimising $d_1(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$. To be precise, $d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$ can be minimised by choosing the major and minor axes to be infinitely long (which makes $d_2(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E) = 0$). The reason for this fallacious result is the lack of a 'scaling' or 'normalizing' factor for the matrix in the criterion. This normalizing factor will counteract the effect of when it becomes too large or too small, thus eliminating this absurd solution to the minimisation problem. While many different scaling factors can be proposed, for reasons described shortly, we choose $\frac{n}{2}\log|\Xi|$ as the scaling factor. Thus our distance of choice becomes:

$$d_3(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E) = \sum_{i=1}^{n} (\mathbf{p}_i - \boldsymbol{\mu})^{\mathrm{T}} \Xi^{-1} (\mathbf{p}_i - \boldsymbol{\mu}) - \frac{n}{2}\log|\Xi| \tag{51}$$

The reason for choosing this particular normalizing factor is that it makes minimisation of the criterion mathematically equivalent to the following well known statistical problem: estimate the mean vector $\boldsymbol{\mu}$ and dispersion matrix Ξ of a bivariate Gaussian distribution, given a sample of n data points. The solution to this problem is given by the corresponding moments of the sample, namely:

$$\bar{\mathbf{p}} = (\bar{x}, \bar{y})$$

and

$$U = \begin{pmatrix} u_{20} & u_{11} \\ u_{11} & u_{02} \end{pmatrix}$$

where $\bar{\mathbf{x}}$ is the centroid as defined in Eq. (49) and

$$u_{20} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{52}$$

$$u_{02} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{53}$$

$$u_{11} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \tag{54}$$

are the second moments of the region. The proof of this result is long and the reader is therefore referred to Theorem 3.2.1 of Anderson (1984) or any other book on multivariate statistical analysis.

This result implies that the best fitting ellipse (in the sense of minimising distance $d_3(\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n, E)$) has a mean vector $\boldsymbol{\mu} = \bar{\mathbf{p}}$ and matrix $\Xi = U$, which are defined by Eqs. (49) and (52)–(54). At this stage, it remains to identify the parameters l_1, l_2 and ϕ that correspond to this closest fit. These can be identified by substituting our solution for Ξ in Eq. (45), i.e.:

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} l_1^2 & 0 \\ 0 & l_2^2 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = U = \begin{pmatrix} u_{20} & u_{11} \\ u_{11} & u_{02} \end{pmatrix} \tag{55}$$

Simplifying the matrix system in Eq. (55) yields the following three equations:

$$l_1^2\cos^2\phi + l_2^2\sin^2\phi = u_{20} \tag{56}$$

$$\left(l_1^2 - l_2^2\right)\sin\phi\cos\phi = u_{11} \tag{57}$$

$$l_1^2\sin^2\phi + l_2^2\cos^2\phi = u_{02} \tag{58}$$

It remains to solve for three unknowns (l_1, l_2 and ϕ) from three equations. After a bit of algebraic and trigonometric manipulation, we get the following solutions to these equations:

$$\phi = \frac{1}{2}\tan^{-1}\left(\frac{2u_{11}}{u_{20} - u_{02}}\right) \tag{59}$$

$$l_1 = \sqrt{\frac{2\left(u_{20} + u_{02} + \sqrt{4u_{11}^2 + (u_{20} - u_{02})^2}\right)}{u_{11}}} \tag{60}$$

$$l_2 = \sqrt{\frac{2\left(u_{20} + u_{02} - \sqrt{4u_{11}^2 + (u_{20} - u_{02})^2}\right)}{u_{11}}} \tag{61}$$

Thus the best fitting ellipse to a region R can be found by calculating the ellipse parameters using the formulae given in Eqs. (49) and (52)–(54).

References

Ailleres, L., Champenois, M., 1994. Refinements to the Fry method (1979) using image processing. Journal of Structural Geology 16, 1327–1330.

Ailleres, L., Champenois, M., Macaudiere, J., Bertrand, J.M., 1995. Use of image analysis in the measurement of finite strain by the normalized Fry method: geological implications for the 'Zone Houillère' (Briançonnais zone, French Alps). Mineralogical Magazine 59, 179–187.

Anderson, T.W., 1984. Multivariate Statistical Analysis, 2nd ed, John Wiley.

Bhattacharyya, P., Hudleston, P., 2001. Strain in ductile shear zones in the Caledonides of northern Sweden: a three-dimensional puzzle. Journal of Structural Geology 23, 1549–1565.

Borradaile, G.J., 1987. Analysis of strained sedimentary fabrics: review and test. Canadian Journal of Earth Sciences 24, 442–455.

Bresser, G., Walter, R., 1999. A new structural model for the SW Irish Variscides: the Variscan front of the NW European Rhenohercynian. Tectonophysics 309, 197–209.

Dunnet, D., 1969. A technique of finite strain analysis using elliptical particles. Tectonophysics 7, 117–136.

Dunnet, D., Siddans, A.W.B., 1971. Non-random sedimentary fabrics and their modification by strain. Tectonophysics 12, 307–325.

Elliott, D., 1970. Determination of finite strain and initial shape from deformed elliptical objects. Geological Society of America Bulletin 81, 2221–2236.

Erslev, E.A., 1988. Normalized center-to-center strain analysis of packed aggregates. Journal of Structural Geology 10, 201–209.

Erslev, E.A., Ge, H., 1990. Least squares center-to-center and mean object ellipse fabric analysis. Journal of Structural Geology 8, 1047–1059.

Fitzgibbon, A., Pilu, M., Fisher, R.B., 1999. Direct least square fitting of ellipses. IEEE Transactions: Pattern Analysis and Machine Intelligence 21, 476–480.

Fry, N., 1979. Random point distributions and strain measurement in rocks. Tectonophysics 60, 806–807.

Gonzalez, R.C., Wintz, P., 1987. Digital Image Processing, Addison-Wesley, Reading, MA.

Halir, R., Flusser, J., 1998. Numerically stable direct least squares fitting of ellipses. In: Proceedings of the 6th International Conference in Central Europe on Computer Graphics and Visualization. WSCG '98. CZ, Plzeò, 1998, pp. 125–132.

Hart, D., Rudman, A.J., 1997. Least-squares fit of an ellipse to anisotropic polar data: application to azimuthal resistivity surveys in karst regions. Computers and Geosciences 23, 189–194.

Heilbronner, R., 2000. Automatic grain boundary detection and grain size analysis using polarization micrographs or orientation images. Journal of Structural Geology 22, 969–981.

Hu, M.K., 1962. Visual pattern recognition using moment invariants. IRE Transactions on Information Theory IT-8, 179–187.

Jähne, B., 1997. Digital Image Processing: Concepts, Algorithms, and Scientific Applications, Springer-Verlag, Berlin.

Jain, A.K., 1989. Fundamentals of Digital Image Processing, Random House, New York.

Lisle, R.J., 1977a. Clastic grain shape and orientation in relation to cleavage from the Aberystwyth Grits, Wales. Tectonophysics 39, 381–395.

Lisle, R.J., 1977b. Estimation of the tectonic strain ratio from the mean shape of deformed elliptical markers. Geologie en Mijnbouw 56, 140–144.

Lisle, R.J., 1994. Paleostrain analysis. In: Hancock, P.L., (Ed.), Continental Deformation, Pergamon Press, pp. 28–42.

McNaught, M., 1994. Modifying the normalized Fry method for aggregates of non-elliptical grains. Journal of Structural Geology 16, 493–503.

Matthews, P.E., Bond, R.A.B., Van Den Berg, J.J., 1974. An algebraic method of strain analysis using elliptical markers. Tectonophysics 24, 31–67.

Mazzoli, S., Zampetti, V., Zuppetta, A., 2001. Very low temperature, natural deformation of fine-grained limestone: a case study from the Lucania region, southern Apennines, Italy. Geodinamica Acta 14, 213–230.

Meere, P.A., 1995. The structural evolution of the western Irish Variscides: an example of obstacle tectonics? Tectonophysics 246, 97–112.

Mulchrone, K.F., 2002. A statistic for estimating strain with confidence intervals from deformed line distributions with an application to schists and gneisses of the Western Gneiss Region, west central Norway. Journal of Structural Geology 24, 545–556.

Mulchrone, K.F., 2003. Application of Delaunay triangulation to the nearest neighbour method of strain analysis. Journal of Structural Geology 25, 689–702.

Mulchrone, K.F., Meere, P.A., 2001. A Windows program for the analysis of tectonic strain using deformed elliptical markers. Computers and Geosciences 27, 1253–1257.

Mulchrone, K.F., O'Sullivan, F., Meere, P.A., 2003. Finite strain estimation using the mean radial length of elliptical objects with confidence intervals. Journal of Structural Geology 25, 529–539.

Paterson, S., Yu, H., 1994. Primary fabric ellipsoids in sandstones: implications for depositional processes and strain analysis. Journal of Structural Geology 16, 505–517.

Ramsay, J.G., 1967. Folding and Fracturing of Rocks, McGraw-Hill, New York.

Robin, P.F., 1977. Determination of geologic strain using randomly oriented strain markers of any shape. Tectonophysics 42, T7–T16.

Russ, J.C., 1999. The Image Processing Handbook, CRC Press, Florida.

Shimamoto, T., Ikeda, Y., 1976. A simple algebraic method for strain estimation from deformed eillipsoidal objects. 1. Basic theory. Tectonophysics 36, 315–337.

Simancas, J.F., Galindo-Zaldívar, J., Azor, A., 2000. Three-dimensional shape and emplacement of the Cardenchosa deformed pluton (Variscan Orogen, southwestern Iberian Massif). Journal of Structural Geology 22, 489–503.

Teague, M.R., 1980. Image analysis via the general theory of moments. Journal of the Optical Society of America 70, 920–930.

Turkowski, K., 1997. Computing 2D polygon moments using Green's theorem. Apple Technical Report. http://www.worldserver.com/turk/computergraphics/Moments.pdf.

Yang, L., Albregtsen, F., 1996. Fast and exact computation of Cartesian geometric moments using discrete Green's theorem. Pattern Recognition 29, 1061–1073.

Zhang, Z., 1997. Parameter estimation techniques: a tutorial with application to conic fitting. Image and Vision Computing 15, 59–76.

Zulauf, G., Kowalczyk, G., Krahl, J., Petschick, R., Schwanz, S., 2002. The tectonometamorphic evolution of high-pressure low-temperature metamorphic rocks of eastern Crete, Greece: constraints from microfabrics, strain, illite crystallinity and paleodifferential stress. Journal of Structural Geology 24, 1805–1828.